

# Psychological Medicine

## Cigarette smoking as a risk factor for schizophrenia or all non-affective psychoses

--Manuscript Draft--

<b>Manuscript Number:</b>	PSM-D-19-00982R1
<b>Full Title:</b>	Cigarette smoking as a risk factor for schizophrenia or all non-affective psychoses
<b>Article Type:</b>	Original Article
<b>Corresponding Author:</b>	Michael King, BSc, MD, PhD University College Medical School London, UNITED KINGDOM
<b>Corresponding Author Secondary Information:</b>	
<b>Corresponding Author's Institution:</b>	University College Medical School
<b>Corresponding Author's Secondary Institution:</b>	
<b>First Author:</b>	Michael King, BSc, MD, PhD
<b>First Author Secondary Information:</b>	
<b>Order of Authors:</b>	Michael King, BSc, MD, PhD Rebecca Jones Irene Petersen Fiona Hamilton Irwin Nazareth
<b>Order of Authors Secondary Information:</b>	
<b>Manuscript Region of Origin:</b>	UNITED KINGDOM
<b>Abstract:</b>	<p><b>Background</b> Smoking tobacco is regarded as an epiphenomenon in patients with schizophrenia when it may be causal. We aimed to examine whether smoking status is related to the onset of schizophrenia or the broader diagnosis of non-affective psychosis, including schizophrenia.</p> <p><b>Methods</b> We used data from The Health Improvement Network primary care database to identify people aged 15 - 24 between 1 January 2004 and 31 December 2009. We followed them until the earliest of: first diagnosis of schizophrenia (or psychosis), patient left the practice, practice left THIN, patient died, or 31 December 2014.</p> <p><b>Results</b> In men, incidence rates for schizophrenia per 100,000 person years at risk were higher in smoking initiators (non-smoker who became a smoker during the study) than in non-smokers (adjusted IRR: 1.94; 95% CI: 1.29 to 2.91) and higher still in smokers (adjusted IRR: 3.32; 95% CI: 2.67 to 4.14). Among women, incidence rate of schizophrenia in smokers was higher than in non-smokers (adjusted IRR: 1.50; 95% CI: 1.06 to 2.12), but no higher than in smoking initiators. For non-affective psychosis, the pattern was similar for men but more evident in women where psychosis incidence rates were higher in smoking initiators (adjusted IRR: 1.90; 95% CI: 1.40 to 2.56) and in smokers (adjusted IRR: 2.13; 95% CI: 1.76 to 2.57) than in non-smokers.</p> <p><b>Conclusions</b> We found an important and strong association between smoking and incidence of schizophrenia. Smoking may increase risk through as yet unknown pathways or smoking may share genetic risk with schizophrenia and non-affective psychoses.</p>

## **Cigarette smoking as a risk factor for schizophrenia or all non-affective psychoses**

**Michael King\*‡**

**Rebecca Jones‡**

**Irene Petersen†**

**Fiona Hamilton†**

**Irwin Nazareth†**

‡Division of Psychiatry

University College London

B Wing, 6th Floor

Maple House

149 Tottenham Court Road

London, W1T 7NF

†Research Department of Primary Care & Population Health

Institute of Epidemiology and Health Care

University College London

Upper 3<sup>rd</sup> Floor, Royal Free Campus

Rowland Hill Street

London NW3 2PF

## **Abstract**

### **Background**

Smoking tobacco is regarded as an epiphenomenon in patients with schizophrenia when it may be causal. We aimed to examine whether smoking status is related to the onset of schizophrenia or the broader diagnosis of non-affective psychosis, including schizophrenia.

### **Methods**

We used data from The Health Improvement Network primary care database to identify people aged 15 - 24 between 1 January 2004 and 31 December 2009. We followed them until the earliest of: first diagnosis of schizophrenia (or psychosis), patient left the practice, practice left THIN, patient died, or 31 December 2014.

### **Results**

In men, incidence rates for schizophrenia per 100,000 person years at risk were higher in smoking initiators (non-smoker who became a smoker during the study) than in non-smokers (adjusted IRR: 1.94; 95% CI: 1.29 to 2.91) and higher still in smokers (adjusted IRR: 3.32; 95% CI: 2.67 to 4.14). Among women, incidence rate of schizophrenia in smokers was higher than in non-smokers (adjusted IRR: 1.50; 95% CI: 1.06 to 2.12), but no higher than in smoking initiators. For non-affective psychosis, the pattern was similar for men but more evident in women where psychosis incidence rates were higher in smoking initiators (adjusted IRR: 1.90; 95% CI: 1.40 to 2.56) and in smokers (adjusted IRR: 2.13; 95% CI: 1.76 to 2.57) than in non-smokers.

### **Conclusions**

We found an important and strong association between smoking and incidence of schizophrenia. Smoking may increase risk through as yet unknown pathways or smoking may share genetic risk with schizophrenia and non-affective psychoses.

## **INTRODUCTION**

People with psychosis are three to six times more likely to smoke than people without(1). Nicotine may relieve stress as well as cognitive symptoms, and counteract the dopamine blockade of antipsychotic medication(2, 3). Smoking may also influence the onset of schizophrenia and other non-affective psychoses, possibly through its impact on neurotransmitter activity in the brain(4, 5). However, distinguishing nicotine's effects from cannabis, an established risk factor for schizophrenia(6, 7) (8), is difficult. We need to know whether smoking tobacco is related to the onset of schizophrenia rather than an associated behaviour once the disorder develops. One way to do this is to use primary care electronic health records in which smoking data, psychiatric diagnoses and a number of other clinical and social variables are recorded over the life course. These clinical data provide a unique resource to test the hypothesis that smoking initiation is associated with an increased risk of developing psychosis.

### **Research question and hypotheses**

#### *Primary research question:*

Is smoking, as recorded in primary care data, related to onset of schizophrenia?

#### *Hypotheses*

1. Smoking initiation is associated with an increased risk of developing non-affective psychosis, including schizophrenia.
2. There is a dose response relationship between smoking and risk of schizophrenia

## **METHODS**

### **Data Source**

We used data from The Health Improvement Network (THIN) primary care database, which contains data on over 12 million patients (including 3.7 million active patients) from over 500 general practices across the UK (<https://www.the-health-improvement-network.co.uk/#thin-research>). There are currently 8732 general practices in the UK. This database contains real time clinical data from GPs' electronic records which include symptoms, diagnosis and lifestyle behaviour such as smoking. The database is broadly representative of UK general practices in terms of patients' age and sex, practice size and geographical distribution. Furthermore, THIN practices contain a profile of patients that is representative of the UK population (9).

#### *Recording of smoking in GP records*

The U.K.'s quality and outcomes framework (QOF) for primary medical care was introduced in 2004 and rewards practices for meeting agreed performance targets(10). Indicators for recording smoking status and providing advice to smokers have featured since QOF was introduced. Thus, smoking status is relatively well recorded in UK electronic primary care records. In 2008-2009 smoking status was recorded for 84% of newly registered patients within 1 year of registration(11).

### **Cohort**

This study included data from all general practices that had contributed data to THIN in accordance with agreed standards. Patients were thus only eligible for entry after the date on which their practice had met the predefined standards for acceptable computer use and acceptable mortality recording (12, 13).

We identified all young people aged from 15 to 24 between 1 January 2004 and 31 December 2009. Individuals were eligible for entry to the cohort from the latest of: date of registration with the practice, dates of acceptable computer use and acceptable mortality recording by the practice, date of their fifteenth birthday, or 1 January 2004.

We followed up individuals until the earliest of: first diagnosis of schizophrenia (or psychosis), patient leaving the practice, practice leaving THIN, patient death, or 31 December 2014. Patients were aged between 15 and 34 years on leaving the study (some were in the cohort for less than a year), enabling us to observe them for up to 10 years, depending upon the date they entered the cohort.

Read codes are a list of clinical terms that have been used by the UK's National Health Service since 1985. They provide a standard vocabulary for clinicians to record patient findings and procedures. Our main exposures were smoking status and smoking intensity, which were defined according to Read codes in medical and additional health data records using established methods to compile the code lists(14). We are aware that individuals may have started smoking earlier, but smoking was defined as starting from the date of first record of smoking status. If someone had a record of smoking they were defined as a smoker; if they had a record of being a non smoker they were defined as a non smoker. People whose status changed from never smoker to smoker during the observation period were classed as smoking initiators. Those who had no record indicating that they were a smoker and no record indicating that they were a non smoker were excluded from the analysis. Patients were also excluded if they had a record of being an ex-smoker before being a non-smoker or a first smoking record, as this indicated misclassification of smoking status. An indicator of smoking intensity (never, light, moderate, heavy) was used based on highest recorded smoking intake during the first 12 months of the observation period.

Smoking intensity was classified by GPs as light (1-9 cigarettes per day), moderate (10-19 cigarettes per day), heavy (20-39 cigarettes per day) or very heavy (more than 40 cigarettes per day) when entering data into the system. Our primary outcome was schizophrenia and our secondary outcome was non-affective psychosis including schizophrenia. Clinicians may be reluctant to give a diagnosis of schizophrenia at first presentation and use a wider or less sensitive term. Thus, we also undertook a secondary analysis in which we included non-affective psychosis in addition to schizophrenia. The disorders were defined according to Read codes indicating a new diagnosis. The code list for a diagnosis of schizophrenia and non-affective psychosis was reviewed by two clinicians to ensure that affective and organic psychotic disorders were excluded. This list is available from the authors. Individuals who only had Read code records which were indicative of pre-existing schizophrenia or repeated episodes (such as chronic paranoid schizophrenia) were excluded. We likewise excluded individuals who had a record of a schizophrenia diagnosis or other non-affective psychotic disorder before their first smoking status record or up to 12 months afterwards. We chose 12 months as the entry of a diagnosis of SCZ often occurs after several months of dysfunction. This helped us protect against reverse causality, because we used smoking records from this period to define smoking intensity. For the main analyses, we therefore started the follow up period at 12 months after entry to the study (see analysis below). The association between smoking and schizophrenia may potentially be confounded by social deprivation and recreational drug use. The Townsend score has been used for decades in the UK as a measure of social deprivation(15). It is derived from variables such as car ownership, owner occupier housing and overcrowded housing. We defined social deprivation in quintiles of the Townsend score for an individual up to 12 months after entry to the cohort. We excluded any GP practices where Townsend scores were poorly recorded.

We defined a binary indicator of history of recreational drug use on the basis of Read codes and once again, we used all records up to 12 months after entry to the cohort in generating this variable.

The study design is illustrated in Figure 1.

### **Analysis**

Binary indicator variables were created to identify patients who had ever had a diagnosis of (1) schizophrenia, or more broadly (2) non-affective psychosis (including schizophrenia).

The earliest recorded date of either schizophrenia or psychotic disorder was taken as the event date for both schizophrenia and psychosis. Incident rates of schizophrenia and non-affective psychosis (including schizophrenia) were estimated as the number of new diagnoses per 100,000 person years at risk (PYAR) from 12 months after entry to the study until end of follow up (defined as the earliest of: date of first diagnosis, date of death, date of patient leaving the practice, date of last data collection from practice or 31 December 2014).

We report unadjusted incidence rates by smoking status (non smoker; smoking initiator; smoker) and by smoking intensity (non, light, moderate and heavy smokers, smoking respectively 0, 1-9, 10-19 and 20 or more cigarettes per day).

Associations between smoking and incidence of schizophrenia and any non-affective psychosis were examined using multilevel Poisson regression models with (log) age as an offset. A random intercept of practice was specified to account for correlations between patients within practices. Results are reported as incident rate ratios (IRR). Similar models were used to examine the association between smoking intensity and incidence of schizophrenia and any non-affective psychotic disorder. In addition, we report IRR adjusted for quintiles of Townsend deprivation score.

Smoking intensity was defined using records up to 12 months after entry to the study. Individuals who were lost to follow up or with a first diagnosis during this period were excluded from the analysis. We did not perform any formal tests for a dose response effect, but rather inferred this relationship from the increasing magnitude of the estimates over the categories of increasing smoking and smoking intensity. A further analysis investigated the sensitivity of results to different methods of defining intensity by using smoking records up to 24 months after patient entry. This was conducted in order to obtain more and better information on smoking intensity.

We explored the possibility that confounding by cannabis use could explain associations between smoking and schizophrenia using the rule out method described by Schneeweiss (2006) (16) (Figure 2). This is a graphical method that quantifies the strength of confounding (i.e. the association between smoking and cannabis use AND the association between cannabis use and schizophrenia) needed to rule out an observed association. The graphs were informed by data on the prevalence of smoking and prevalence of cannabis as well as the IRR and 95%CI for the associations between smoking and schizophrenia. We did this using a combination of results from our study and figures reported in the literature.

Thus, we estimated prevalence of smoking from our own data of 31.1% for men and 27.8% for women, and an estimated prevalence of cannabis use of 20.4% for men and 12.8% for women from the 2016-2017 crime survey for England and Wales which is representative of the UK population(17). These figures were appropriate, as the THIN population is representative of the UK population. The resulting curves divide the graph area in two. The area to the right and above each curve represents when an apparent association could be explained by confounding, i.e. combinations of associations with cannabis use (the potential confounder) which would reduce the IRR or the lower bound of the 95% CI to 1. In contrast,

the area to the left and below each curve represents the situations when association between smoking and schizophrenia/psychosis cannot be ruled out by residual confounding. In Figure 2, we produced separate graphs for men and women and for incidence of schizophrenia and non-affective psychosis and interpreted them in conjunction with associations with cannabis use reported in the literature.

The THIN scheme for obtaining and providing anonymous patient data to researchers was approved by the NHS South-East Multicentre Research Ethics Committee in 2002. Approval for this study was obtained from the Scientific Review Committee (protocol reference number 16THIN041)

## **RESULTS**

### *Identification of cohort*

Initially, data on 1,056,176 patients were extracted. Smoking records from within the first 12 months after entry were used to establish smoking status and intensity. The observation period for incidence of schizophrenia or non-affective psychotic disorder thus started 12 months after entry to the study. 907,586 patients were retained in the analysis data set after excluding those who were lost to follow up before the start of the observation period (n=104,937), those with a first diagnosis earlier than 1 year after the date of first record of smoking (n=392), those missing a Townsend deprivation score (n=42,990) and patients whose smoking status could not be ascertained from Read codes (n=271).

### *Description of cohort*

*Diagnosis:* Of the 907,586 included in the analysis 536 (0.06%) received a diagnosis of schizophrenia and 1,582 (0.17%) received a diagnosis of non-affective psychotic disorder

(including schizophrenia). Overall, 258,170 (28.5%) of the cohort were aged 15 years on entry to the study, 267,771 (29.5%) aged 16 to 20 years, 307,256 (33.9%) aged 21 to 25 years and 74,389 (8.2%) aged 26 to 30 years. There were slightly more women (52.6%) than men (47.4%). The mean age of first diagnosis with a non-affective psychotic disorder (including schizophrenia) was 24.0 years (SD: 4.1) in males and 24.5 years (SD: 4.5) in females. The median length of follow up was 6.0 years (Inter Quartile Range: 3.4 to 9.0 years).

*Smoking:* 570,457 (62.9%) were classified as non-smokers throughout the whole observation period, 266,396 (29.4%) had started smoking prior to the observation period (smokers) and 70,733 (7.8%) started smoking during follow up (smoking initiators). The mean age of first record of smoking was 21.1 years (SD: 3.8) in males and 20.0 years (SD: 3.6) in females.

Based on smoking records up to 12 months after entry to the study, 465,192 (51.3%) patients were classified as never smokers, 71,503 (7.9%) as light ( $\leq 9$  cigarettes per day), 86,473 (9.5%) as moderate (10-19 cigarettes per day) and 31,930 (3.5%) as heavy smokers ( $\geq 20$  cigarettes per day). It was not possible to determine the level of smoking for 252,488 young people (27.8%) and these individuals were therefore not included in analyses of smoking intensity.

Using smoking records up to 24 months after entry to the study, 520,421 (57.32%) of patients were classified as never smokers, 80,332 (8.85%) as light, 100,631 (11.09%) as moderate and 39,494 (4.35%) as heavy smokers. It was not possible to determine level of smoking for 166,888 (18.39%) of patients.

*Recreational drug use:* In total, 33,512 (3.7%) individuals had a record indicating a history of recreational drug use. This is much lower than other estimates of recreational drug use

made in other epidemiological studies suggesting that many individuals who had used recreational drugs did not have a record of this in their primary care data and we therefore decided not to adjust for this covariate in our analysis.

*Incidence rates by smoking status and intensity*

*Incidence rates for schizophrenia:* Among men, the incidence rates for schizophrenia per 100,000 PYAR were higher in smoking initiators than in non-smokers (adjusted IRR: 1.94; 95% CI: 1.29 to 2.91) and higher still in smokers (adjusted IRR: 3.32; 95% CI: 2.67 to 4.14).

There was evidence of a dose response effect for amount of smoking, with an effect size of markedly greater magnitude in heavy smokers (adjusted IRR: 5.54; 95% CI: 3.84 to 7.99) than in moderate smokers (adjusted IRR: 3.80; 95% CI: 2.74 to 5.27). Adjustment for deprivation reduced the magnitude of the IRRs very slightly, suggesting only slight confounding (Table 1).

Among women, there was evidence of an elevated incidence rate of schizophrenia in smokers compared to non-smokers (adjusted IRR: 1.50; 95% CI: 1.06 to 2.12), but not in individuals who started smoking during the observation period. Similarly, only heavy smokers showed evidence of elevated rates of schizophrenia incidence compared to non-smokers (adjusted IRR: 3.74; 95% CI: 2.10 to 6.64). Again our adjustments had little impact on the result (Table 1).

*Incidence rates for non-affective psychosis (including schizophrenia):* Among men, psychosis incidence rates were higher in smoking initiators than in non-smokers (adjusted IRR: 2.65; 95% CI: 2.15 to 3.27) and higher still in smokers (adjusted IRR: 2.96; 95% CI: 2.59 to 3.38).

There was evidence of a dose response effect for amount of smoking, with effect sizes of increasing magnitude in light (adjusted IRR: 2.72; 95% CI: 2.18 to 3.39), moderate (adjusted

IRR: 3.22; 95% CI: 2.67 to 3.88) and heavy smokers (adjusted IRR: 4.07; 95% CI: 3.26 to 5.09) relative to non-smokers (Table 2).

Among women, psychosis incidence rates were higher in smoking initiators than in non-smokers (adjusted IRR: 1.90; 95% CI: 1.40 to 2.56) and also in smokers (adjusted IRR: 2.13; 95% CI: 1.76 to 2.57). There was likewise evidence of a dose response effect for amount of smoking, with effect sizes of increasing magnitude in light (adjusted IRR: 2.20; 95% CI: 1.65 to 2.91), moderate (adjusted IRR: 2.31; 95% CI: 1.78 to 3.01) and heavy smokers (adjusted IRR: 4.10; 95% CI: 2.97 to 5.66) relative to non-smokers (Table 2).

*Sensitivity analysis:* The impact of smoking intensity was also examined based on smoking records up to 24 months after the start of the observation period (Table 3). This was to allow for a longer period of prodromal symptoms as might be expected before an entry of a diagnosis of schizophrenia or non-affective psychosis was made in the records. Higher smoking intensity in the 24 months before diagnosis is strongly linked in both sexes to an increased likelihood of either schizophrenia or the more inclusive category of non-affective psychosis.

*Sensitivity analysis examining confounding by cannabis use:* Based on the information available from our own data and an estimated prevalence of cannabis use in the population, we were able to quantify the likelihood of confounding by cannabis use. The graphs for men in Figure 2 suggest that unmeasured confounding needs to be substantial to rule out the observed association between smoking and schizophrenia/psychosis. For example, this would be the case if male smokers were 15 times more likely to also use cannabis than non smokers AND cannabis users were 10 times more likely to develop schizophrenia. Using evidence from the literature, the point represented by an OR of 3.90 for schizophrenia in cannabis users compared to non-users (18) and an OR of 4.13 for cannabis use in smokers

compared to non-smokers (19), falls below and to the left both of the curve for the IRR and the lower bound of its 95% CI. This suggests that cannabis use cannot fully explain the association we found between smoking and schizophrenia. The same is true for the relationship between smoking and psychosis in both men and women. For schizophrenia in women (Figure 2), the point represented by these two ORs falls below the curve for the IRR, but above the curve for its 95% CI, indicating that the data are compatible with confounding by cannabis use.

## DISCUSSION

### *Findings*

Our results show associations between cigarette smoking and incidence of schizophrenia, as well as non-affective psychoses including schizophrenia, after adjustment for social deprivation. In men, incidence of schizophrenia was nearly twice as high in smoking initiators as non-smokers, and three times higher in pre-existing smokers, with a dose response effect of amount of smoking. This pattern was similar, but less pronounced, in women with evidence for a higher incidence rate in smokers compared to non-smokers, but not for smoking initiators.

In men, the findings for non-affective psychosis (including schizophrenia) were similar to those for schizophrenia. However, for women the pattern was more marked than in schizophrenia in that the association held for smoking initiators and smokers, and a dose response relationship occurred across all three categories of smoker. The greater magnitude of the association / dose response effect in non-affective psychosis in women may have occurred because of the low numbers of women with a diagnosis of schizophrenia, resulting in lower power to detect an association in that comparison. The gender differences observed may be an artefact of the generally later onset of schizophrenia in women and the cohort age group.

Our sensitivity analysis revealed that whether smoking intensity was defined using smoking records up to 12 or 24 months after the start of the observation period had little impact on the findings.

Our second sensitivity analysis using the rule out method suggests that it is unlikely that cannabis use can solely explain the association we found between smoking and schizophrenia in men, or between smoking and psychosis in either sex. It is possible that

cannabis use could confound the association between smoking and schizophrenia in women, however very few women were diagnosed with schizophrenia and this rather than a true association may explain this finding.

### *The context of other research*

Our incidence rates for schizophrenia are comparable to the recent overall figure (also from the THIN database) of 9.2 per 100,000 PYAR for people aged 16-65 in the UK(20). Our findings echo the results of other recent studies in this area. In a recent meta-analysis of observational research, Gurillo et al reported a relative risk of new psychotic disorders in daily smokers versus non-smokers of 2.18 (95% CI 1.23–3.85)(4). Daily smokers developed psychotic illness at an earlier age than non-smokers (weighted mean difference –1.04 years, CI –1.82 to –0.26). There is also a suggestion that people with psychosis may start smoking at an earlier age than healthy controls (–0.44 years, CI –1.21 to 0.34). An association between early age of initiation of smoking and non-affective psychosis has also been reported in an Australian birth cohort study(21). Similar (albeit small) associations have been found between use of nicotine in smokeless form and onset of schizophrenia(22). Although there are cultural differences in that prevalence of smoking in patients with schizophrenia in non-western countries is lower (at least in women)(23-25), a higher risk in psychotic people persists. In a review of Chinese data(24), prevalence of smoking in women with schizophrenia was 4.3% compared with a general population rate of 3.4%. Although not reported in the publication, this amounts to an odds ratio (OR) of 1.28 for smoking in women with schizophrenia.

### *Strengths and Limitations*

This observational study has several strengths in taking account of time from exposure to outcome, magnitude and direction of the association, biological gradient (intensity of

smoking), consistency, and specificity. Most important of all our results are generalisable as the analysis was conducted in a very large sample that was representative of the UK population and was longitudinal in nature. Nevertheless, there are limited demographic variables in THIN for which adjustment is possible and we cannot rule out residual confounding. In particular, we cannot be certain whether smoking patterns disguised a potential effect from recreational drugs, especially cannabis(26). The prevalence of data on recreational drug use was considered too low to be plausible in this age group. The 2016-2017 crime survey for England and Wales found that 19% of young people aged 16 to 24 had taken a drug in the previous year (and the figure is even higher a decade earlier)(27). Thus, primary care records have failed to identify 15% of the young people in our cohort, who have probably used an illegal drug.

We are conscious that we cannot rule out smoking initiation as a part of the prodrome of psychosis, given that prodromal symptoms (social isolation, academic difficulties, odd behaviours, and symptoms of depression or anxiety) may precede the appearance of frank schizophrenia by an average of 23 months(28). Our stipulation of at least a year between entry to the cohort and onset of schizophrenia, however, may go some way towards excluding this possibility. Our mean age of 'first smoking' may be slightly higher than other estimates, although it depends on how this is defined. Recent estimates suggest the most common age of *first trying* a cigarette is 10–15 years(29), while take up of regular smoking usually continues up to the early 20s(30). If participants with multiple smoking records had a record of being a smoker followed by a record of being a non smoker or an ex smoker, we inferred that they had given up smoking. However, individuals may have stopped smoking but lacked any record of having done so (i.e. becoming an ex or non smoker). Finally, smoking was not necessarily recorded during every appointment and, even if it had been,

may not have captured every change in an individual's smoking history. Thus, a degree of misclassification of smoking status and intensity is likely in our dataset.

### *Genetic factors*

There is a debate about possible genetic explanations for our observations. For example, it has been suggested that nicotine may overcome an observed dysregulation in schizophrenia in the expression of genes that encode the nicotinic acetylcholine receptor (nAChR) “ $\alpha 7$ ” (31, 32). However, an unbalanced gene expression could precede the onset of psychotic symptoms, with nicotine abuse and dependence merely acting as indicators of future psychosis (33). Another polymorphism (*rs1051730*) in the nicotinic acetylcholine receptor gene cluster *CHRNA5-A3-B4* has been reported to be associated with schizophrenia *and* with smoking intensity, but not with smoking initiation or with depression. A further study conducted in Iceland also supports the notion of pleiotropy for the observed comorbidity of addiction (including nicotine) and psychosis (34). In another approach to the study of possible causal effects of smoking on schizophrenia, Gage and colleagues reported little evidence of a causal association between smoking initiation and schizophrenia in a study using Mendelian randomisation (35). However, they could not exclude the possibility of a causal effect of heavier, lifetime exposure, rather than initiation, on the disorder. A more recent study that also used Mendelian randomisation in the UK Biobank data has indicated that the association between smoking and schizophrenia may be due in part to a causal effect of smoking (36).

### *Interpretation*

The magnitude and strength of the association between smoking and incidence of schizophrenia supports suggestions that smoking involves a risk through as yet unknown

pathways, or that it is an addictive condition which shares genetic risk with schizophrenia and other non- psychotic psychoses (34).

**ACKNOWLEDGEMENTS**

MK and RJ acknowledge the support of the UCLH NIHR Biomedical Research Centre.

**CONFLICT OF INTEREST**

There are no conflicts of interest for any of the authors

**FUNDING**

No independent funding was obtained to undertake this study.

## REFERENCES

1. de Leon J, Diaz FJ. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophrenia Research*. 2005;76(2–3):135-57.
2. Winterer G. Why do patients with schizophrenia smoke? *Current Opinion in Psychiatry*. 2010;23(2):112-9.
3. Whitton A, Green A, Pizzagalli D, Williams J, Brunette M. S244. Potent Dopamine D2 Receptor Antagonists Block the Reward-Enhancing Effects of Nicotine in Smokers With Schizophrenia. *Biological Psychiatry*. 2018;83(9, Supplement):S443.
4. Gurillo P, Jauhar S, Murray RM, MacCabe JH. Does tobacco use cause psychosis? Systematic review and meta-analysis. *The Lancet Psychiatry*. 2015;2(8):718-25.
5. Large M, MacCabe JH. Tobacco and psychosis: Not quite a smoking gun. *Australian and New Zealand Journal of Psychiatry*. 2015;49(10):859-60.
6. Arseneault L, Cannon M, Wittton J, Murray RM. Causal association between cannabis and psychosis: examination of the evidence. *The British Journal of Psychiatry*. 2004;184(2):110-7.
7. Gage SH, Hickman M, Zammit S. Association Between Cannabis and Psychosis: Epidemiologic Evidence. *Biological Psychiatry*. 2016;79(7):549-56.
8. Murray RM, Bhavsar V, Tripoli G, Howes O. 30 Years on: How the Neurodevelopmental Hypothesis of Schizophrenia Morphed into the Developmental Risk Factor Model of Psychosis. *Schizophrenia Bulletin*. 2017:sbx121-sbx.
9. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. *Journal of Innovation in Health Informatics*. 2011;19(4):251-5.
10. Taggar JS, Coleman T, Lewis S, Szatkowski L. The impact of the Quality and Outcomes Framework (QOF) on the recording of smoking targets in primary care medical records: cross-sectional analyses from The Health Improvement Network (THIN) database. *BMC Public Health*. 2012;12(1):1-11.
11. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, White IR, et al. Smoker, ex-smoker or non-smoker? The validity of routinely recorded smoking status in UK primary care: a cross-sectional study. *BMJ Open*. 2014;4(4).
12. Horsfall L, Walters K, Petersen I. Identifying periods of acceptable computer usage in primary care research databases. *Pharmacoepidemiology and Drug Safety*. 2013;22(1):64-9.
13. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. *Pharmacoepidemiology and Drug Safety*. 2009;18(1):76-83.
14. Dave S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. *PharmacoepidemiolDrug Saf*. 2009;18(8):704-7.
15. Dolan SA, Jarman B, Bajekal M, Davies PM, Hart D. Measuring disadvantage: changes in the underprivileged area, Townsend, and Carstairs scores 1981-91. *Journal of Epidemiology and Community Health*. 1995;49(Suppl 2):S30-S3.
16. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and Drug Safety*. 2006;15(5):291-303.
17. Office H. Drug misuse: findings from the 2017 to 2018 CSEW. Appendix table 3.02. 2018.
18. Marconi A, Di Forti M, Lewis CM, Murray RM, Vassos E. Meta-analysis of the Association Between the Level of Cannabis Use and Risk of Psychosis. *Schizophrenia Bulletin*. 2016;42(5):1262-9.
19. Fix BV, Smith D, O'Connor R, Heckman BW, Willemsen MC, Cummings M, et al. Cannabis use among a nationally representative cross-sectional sample of smokers and non-smokers in the

- Netherlands: results from the 2015 ITC Netherlands Gold Magic Survey. *BMJ Open*. 2019;9(3):E024497.
20. Hardoon S, Hayes JF, Blackburn R, Petersen I, Walters K, Nazareth I, et al. Recording of severe mental illness in United Kingdom primary care, 2000-2010. *PloS one*. 2013;8(12):e82365-e.
  21. McGrath JJ, Alati R, Clavarino A, Williams GM, Bor W, Najman JM, et al. Age at first tobacco use and risk of subsequent psychosis-related outcomes: A birth cohort study. *Australian and New Zealand Journal of Psychiatry*. 2015.
  22. Munafò MR, Larsson Lönn S, Sundquist J, Sundquist K, Kendler K. Snus use and risk of schizophrenia and non-affective psychosis. *Drug and Alcohol Dependence*. 2016;164:179-82.
  23. Aguocha CM, Aguocha JK, Igwe M, Uwakwe RU, Onyeama GM. Prevalence and correlates of cigarette smoking among patients with schizophrenia in southeast Nigeria. *Acta Psychiatrica Scandinavica*. 2015;131(3):206-12.
  24. Cao X-L, Li Y, Zhong B-L, Chiu HFK, Ungvari GS, Ng CH, et al. Current cigarette smoking in Chinese female patients with schizophrenia: A meta-analysis. *Psychiatry Research*. 2016;235:203-5.
  25. Li Y, Hou C-L, Ma X-R, Zang Y, Jia F-J, Zhong B-L, et al. Smoking and its associations with sociodemographic and clinical characteristics and quality of life in patients with schizophrenia treated in primary care in China. *General Hospital Psychiatry*. 2016;38:79-83.
  26. Fergusson DM, Hall W, Boden JM, Horwood LJ. Rethinking cigarette smoking, cannabis use, and psychosis. *The Lancet Psychiatry*. 2015;2(7):581-2.
  27. Broadfield D. Drug Misuse: Findings from the 2016/2017 Crime Survey for England and Wales. In: Office H, editor. *Statistical Bulletin 11/17* ed. UK: Home Office; 2017. p. 2.
  28. Kafali HY, Bildik T, Bora E, Yuncu Z, Erermis HS. Distinguishing prodromal stage of bipolar disorder and early onset schizophrenia spectrum disorders during adolescence. *Psychiatry Research*. 2019;275:315-25.
  29. West R. Tobacco smoking: Health impact, prevalence, correlates and interventions. *Psychology & Health*. 2017;32(8):1018-36.
  30. Dierker L, He J, Kalaydjian A, Swendsen J, Degenhardt L, Glantz M, et al. The importance of timing of transitions for risk of regular smoking and nicotine dependence. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine*. 2008;36(1):87-92.
  31. Kunii Y, Zhang W, Xu Q, Hyde TM, McFadden W, Shin JH, et al. CHRNA7 and CHRFB7A mRNAs: Co-Localized and Their Expression Levels Altered in the Postmortem Dorsolateral Prefrontal Cortex in Major Psychiatric Disorders. *American Journal of Psychiatry*. 2015;172(11):1122-30.
  32. Liu X-a, Kenny PJ.  $\alpha 5$  nicotinic receptors link smoking to schizophrenia. *Nature Medicine*. 2017;23:277.
  33. Akbarian S, Kundakovic M. CHRNA7 and CHRFB7A: Psychosis and Smoking? Blame the Neighbors! *American Journal of Psychiatry*. 2015;172(11):1054-6.
  34. Reginsson GW, Ingason A, Euesden J, Bjornsdottir G, Olafsson S, Sigurdsson E, et al. Polygenic risk scores for schizophrenia and bipolar disorder associate with addiction. *Addiction Biology*. 2018;23(1):485-92.
  35. Gage SH, Jones HJ, Taylor AE, Burgess S, Zammit S, Munafò MR. Investigating causality in associations between smoking initiation and schizophrenia using Mendelian randomization. *Scientific Reports*. 2017;7:40653.
  36. Wootton RE, Richmond RC, Stuijzand BG, Lawn RB, Sallis HM, Taylor GMJ, et al. Causal effects of lifetime smoking on risk for depression and schizophrenia: Evidence from a Mendelian randomisation study. *bioRxiv*. 2018:381301.

**Table 1: Incidence of schizophrenia by smoking status and intensity**

	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>
<b>Smoking status</b>										
	<b>Men (N = 429,922)</b>					<b>Women (N = 477,664)</b>				
Non smoker	123	14.36	8.6 (7.2 to 10.2)	1	1	76	14.20	5.4 (4.3 to 6.7)	1	1
Smoking initiator	29	2.13	13.6 (9.5 to 19.6)	2.00 (1.33 to 2.99)	1.94 (1.29 to 2.91)	12	2.35	5.1 (2.9 to 9.0)	1.28 (0.69 to 2.35)	1.25 (0.68 to 2.29)
Smoker	239	7.57	31.6 (27.8 to 35.8)	3.56 (2.87 to 4.43)	3.32 (2.67 to 4.14)	57	6.59	8.6 (6.7 to 11.2)	1.65 (1.17 to 2.32)	1.50 (1.06 to 2.12)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> =0.0172	<i>p</i> =0.0711
<b>Smoking intensity *</b>										
	<b>Men (N = 279,007)</b>					<b>Women (N = 376,091)</b>				
Non smoker	81	9.63	8.4 (6.8 to 10.5)	1	1	56	11.69	4.8 (3.7 to 6.2)	1	1
Light smoker	49	1.30	37.6 (28.4 to 49.8)	4.08 (2.86 to 5.82)	3.91 (2.74 to 5.58)	12	2.00	6.0 (3.4 to 10.6)	1.25 (0.67 to 2.33)	1.20 (0.64 to 2.23)
Moderate smoker	67	1.80	37.3 (29.4 to 47.4)	4.04 (2.92 to 5.59)	3.80 (2.74 to 5.27)	19	2.38	8.0 (5.1 to 12.5)	1.73 (1.03 to 2.91)	1.58 (0.94 to 2.67)
Heavy smoker	46	0.76	60.2 (45.1 to 80.3)	6.08 (4.23 to 8.74)	5.54 (3.84 to 7.99)	15	0.73	20.6 (12.4 to 34.2)	4.29 (2.42 to 7.59)	3.74 (2.10 to 6.64)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> <0.0001	<i>p</i> <0.0001

**Notes:** N = number of incident cases, PYAR = person years at risk, IR = incident rate per 100,000 PYAR, IRR = incident rate ratio, CI = confidence interval. Adjusted IRR control for Townsend deprivation score. \* Smoking intensity was defined using records up to 12 months after entry to the study. Light smoker: <=9 cigarettes per day; moderate smoker: 10-19 cigarettes per day; heavy smoker: >=20 cigarettes per day.

**Table 2: Incidence of all non affective psychosis (including schizophrenia) by smoking status and intensity**

	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>
<b>Smoking status</b>										
	<b>Men (N = 429,922)</b>					<b>Women (N = 477,664)</b>				
Non smoker	354	14.36	24.6 (22.2 to 27.4)	1	1	219	14.20	15.4 (13.5 to 17.6)	1	1
Smoking initiator	116	2.13	54.4 (45.4 to 65.3)	2.71 (2.20 to 3.35)	2.65 (2.15 to 3.27)	53	2.35	22.6 (17.3 to 29.6)	1.93 (1.43 to 2.61)	1.90 (1.40 to 2.56)
Smoker	613	7.57	81.0 (74.8 to 87.7)	3.15 (2.76 to 3.59)	2.96 (2.59 to 3.38)	227	6.59	34.4 (30.2 to 39.2)	2.27 (1.88 to 2.73)	2.13 (1.76 to 2.57)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> <0.0001	<i>p</i> <0.0001
<b>Smoking intensity *</b>										
	<b>Men (N = 279,007)</b>					<b>Women (N = 376,091)</b>				
Non smoker	269	9.63	27.9 (24.8 to 31.5)	1	1	173	11.69	14.8 (12.8 to 17.2)	1	1
Light smoker	113	1.30	86.7 (72.1 to 104.3)	2.82 (2.26 to 3.52)	2.72 (2.18 to 3.39)	67	2.00	33.5 (26.3 to 42.5)	2.26 (1.70 to 2.99)	2.20 (1.65 to 2.91)
Moderate smoker	190	1.80	105.8 (91.8 to 122.0)	3.41 (2.83 to 4.11)	3.22 (2.67 to 3.88)	84	2.38	35.3 (28.5 to 43.7)	2.45 (1.89 to 3.18)	2.31 (1.78 to 3.01)
Heavy smoker	114	0.76	149.1 (124.1 to 179.1)	4.45 (3.57 to 5.55)	4.07 (3.26 to 5.09)	49	0.73	67.3 (50.9 to 89.1)	4.47 (3.25 to 6.15)	4.10 (2.97 to 5.66)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> <0.0001	<i>p</i> <0.0001

**Notes:** N = number of incident cases, PYAR = person years at risk, IR = incident rate per 100,000 PYAR, IRR = incident rate ratio, CI = confidence interval. Adjusted IRR control for Townsend deprivation score. \* Smoking intensity was defined using records up to 12 months after entry to the study. Light smoker: <=9 cigarettes per day; moderate smoker: 10-19 cigarettes per day; heavy smoker: >=20 cigarettes per day.

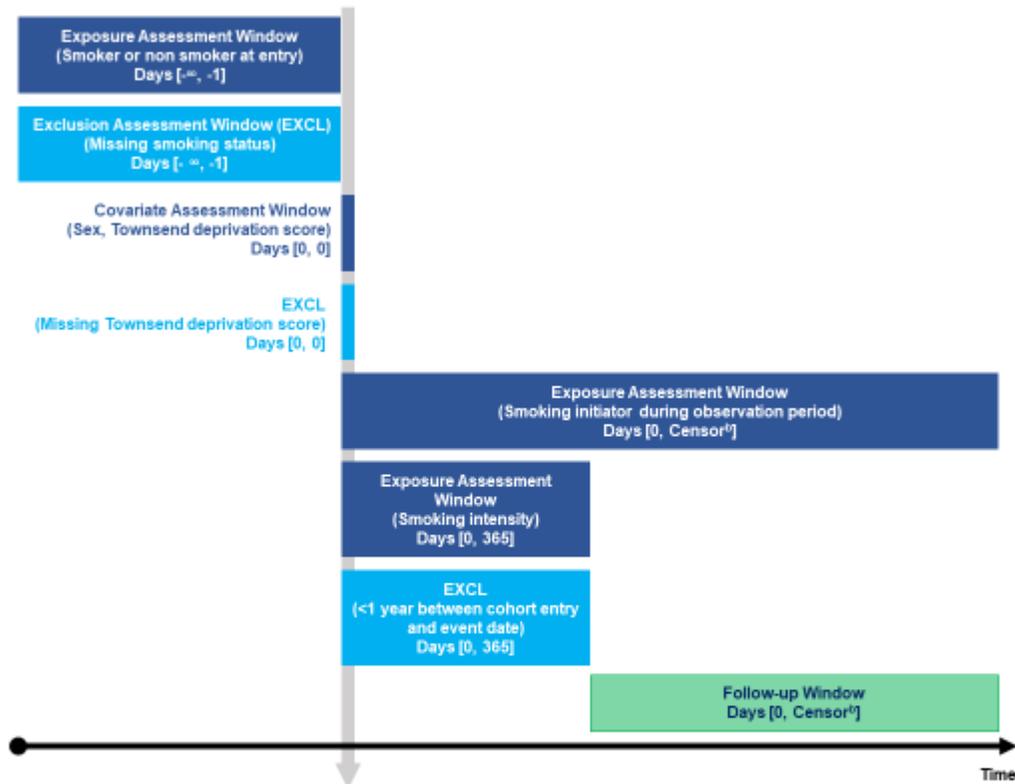
**Table 3: Association between smoking intensity and incidence of schizophrenia or non-affective psychosis based on smoking records up to 24 months after entry to the study**

	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>
<b>Schizophrenia</b>										
	<b>Men (N = 285,562)</b>					<b>Women (N = 355,139)</b>				
Non smoker	80	11.36	7.0 (5.7 to 8.8)	1	1	54	13.04	4.1 (3.2 to 5.4)	1	1
Light smoker	48	1.55	31.0 (23.3 to 41.1)	4.17 (2.91 to 5.96)	4.03 (2.81 to 5.77)	8	2.24	3.6 (1.8 to 7.1)	0.87 (0.41 to 1.82)	0.83 (0.40 to 1.75)
Moderate smoker	66	2.19	30.2 (23.7 to 38.4)	4.06 (2.93 to 5.62)	3.85 (2.78 to 5.35)	14	2.80	5.0 (3.0 to 8.4)	1.25 (0.69 to 2.25)	1.14 (0.63 to 2.05)
Heavy smoker	42	0.97	43.3 (32.0 to 58.6)	5.53 (3.80 to 8.04)	5.09 (3.49 to 7.43)	14	0.93	15.0 (8.9 to 25.3)	3.63 (2.02 to 6.54)	3.15 (1.74 to 5.71)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> =0.0003	<i>p</i> =0.0013
<b>All non affective psychosis</b>										
	<b>Men (N = 285,562)</b>					<b>Women (N = 355,139)</b>				
Non smoker	266	11.36	23.4 (20.8 to 26.4)	1	1	161	13.04	12.3 (10.6 to 14.4)	1	1
Light smoker	106	1.55	68.4 (56.5 to 82.7)	2.77 (2.21 to 3.47)	2.67 (2.13 to 3.34)	50	2.24	22.3 (16.9 to 29.4)	1.81 (1.32 to 2.49)	1.76 (1.28 to 2.42)
Moderate smoker	196	2.19	89.6 (77.9 to 103.1)	3.58 (2.98 to 4.31)	3.39 (2.81 to 4.08)	72	2.80	25.7 (20.4 to 32.4)	2.14 (1.62 to 2.83)	2.01 (1.52 to 2.66)
Heavy smoker	115	0.97	118.6 (98.8 to 142.3)	4.46 (3.58 to 5.55)	4.09 (3.28 to 5.11)	45	0.93	48.2 (36.0 to 64.5)	3.88 (2.78 to 5.40)	3.53 (2.53 to 4.93)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> <0.0001	<i>p</i> <0.0001

**Notes:** N = number of incident cases, PYAR = person years at risk, IR = incident rate per 100,000 PYAR, IRR = incident rate ratio, CI = confidence interval. Adjusted IRR control for Townsend deprivation score. Light smoker: <=9 cigarettes per day; moderate smoker: 10-19 cigarettes per day; heavy smoker: >=20 cigarettes per day.

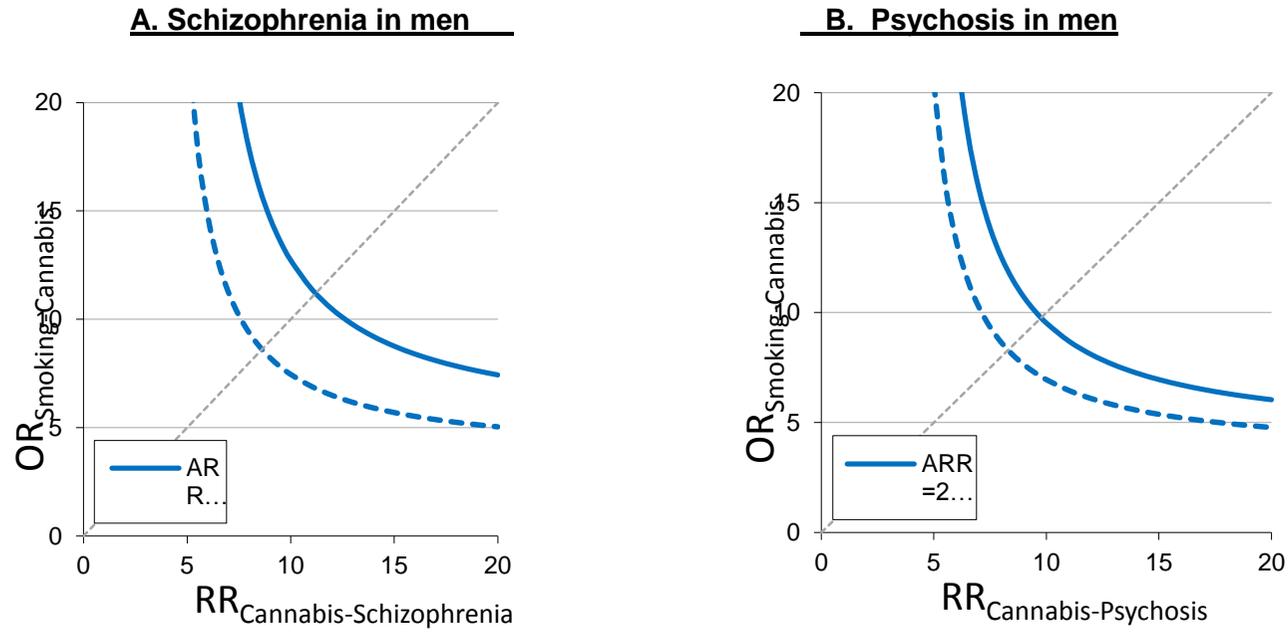
Figure 1 Study Design

Cohort Entry Date (Age 15 to 24 years between 1 January 2004 and 31 December 2009<sup>a</sup>) Day 0

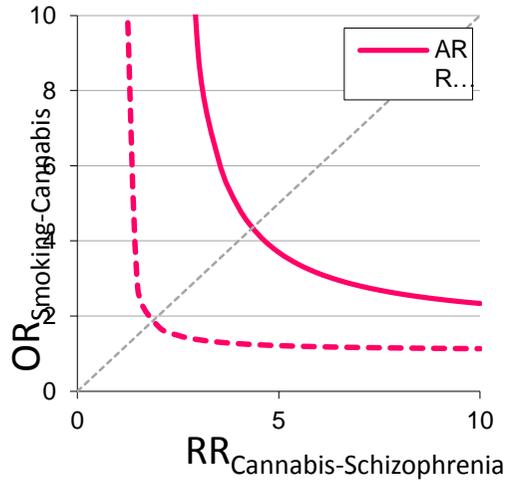


- a. Individuals were eligible for entry to the cohort from the latest of: date of registration with the practice, dates of acceptable computer use and acceptable mortality recording by the practice, date of their fifteenth birthday, or 1 January 2004.
- b. Individuals were followed up until the earliest of: first diagnosis of schizophrenia (or psychosis), patient leaving the practice, practice leaving THIN, patient death, or 31 December 2014.

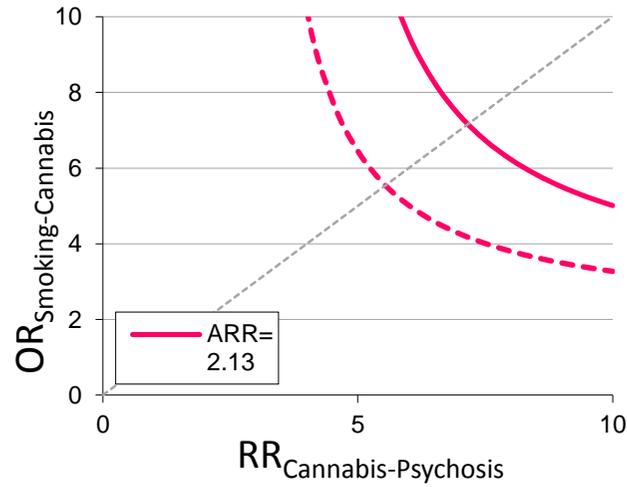
**Figure 2: Cannabis use as a potential confounder of the apparent association between smoking and schizophrenia and non-affective psychosis**



**C. Schizophrenia in women**



**D. Psychosis in women**



**Figure note/caption:** Solid lines represent the adjusted incident rate ratio (IRR) for incidence of schizophrenia or psychosis respectively in smokers compared to non smokers. Dashed lines represent the lower bound of the 95% confidence interval (CI) for the IRR. The areas of the graphs above and to the right of these curves indicate combinations of associations with cannabis use (the potential confounder) which would reduce the IRR or the lower bound of the 95% CI to 1, indicating that the confounding by cannabis use eliminates the association between smoking and schizophrenia or psychosis. Conversely, areas below and to the left of the curves indicate combinations which may result in partial confounding, but where cannabis use cannot fully explain the observed association.

## **Cigarette smoking as a risk factor for schizophrenia or all non-affective psychoses**

**Michael King\*‡**

**Rebecca Jones‡**

**Irene Petersen†**

**Fiona Hamilton†**

**Irwin Nazareth†**

‡Division of Psychiatry

University College London

B Wing, 6th Floor

Maple House

149 Tottenham Court Road

London, W1T 7NF

†Research Department of Primary Care & Population Health

Institute of Epidemiology and Health Care

University College London

Upper 3<sup>rd</sup> Floor, Royal Free Campus

Rowland Hill Street

London NW3 2PF

## Abstract

### Background

Smoking tobacco is regarded as an epiphenomenon in patients with schizophrenia when it may be causal. We aimed to examine whether smoking status is related to the onset of schizophrenia or the broader diagnosis of non-affective psychosis, including schizophrenia.

### Methods

We used data from The Health Improvement Network primary care database to identify people aged 15 - 24 between 1 January 2004 and 31 December 2009. We followed them until the earliest of: first diagnosis of schizophrenia (or psychosis), patient left the practice, practice left THIN, patient died, or 31 December 2014.

### Results

In men, incidence rates for schizophrenia per 100,000 person years at risk were higher in smoking initiators (non-smoker who became a smoker during the study) than in non-smokers (adjusted IRR: 1.94; 95% CI: 1.29 to 2.91) and higher still in smokers (adjusted IRR: 3.32; 95% CI: 2.67 to 4.14). Among women, incidence rate of schizophrenia in smokers was higher than in non-smokers (adjusted IRR: 1.50; 95% CI: 1.06 to 2.12), but no higher than in smoking initiators. For non-affective psychosis, the pattern was similar for men but more evident in women where psychosis incidence rates were higher in smoking initiators (adjusted IRR: 1.90; 95% CI: 1.40 to 2.56) and in smokers (adjusted IRR: 2.13; 95% CI: 1.76 to 2.57) than in non-smokers.

### Conclusions

We found an important and strong association between smoking and incidence of schizophrenia. Smoking may increase risk through as yet unknown pathways or smoking may share genetic risk with schizophrenia and non-affective psychoses.

## INTRODUCTION

People with psychosis are three to six times more likely to smoke than people without(1). Nicotine may relieve stress as well as cognitive symptoms, and counteract the dopamine blockade of antipsychotic medication(2, 3). Smoking may also influence the onset of schizophrenia and other non-affective psychoses, possibly through its impact on neurotransmitter activity in the brain(4, 5). However, distinguishing nicotine's effects from cannabis, an established risk factor for schizophrenia(6, 7) (8), is difficult. We need to know whether smoking tobacco is related to the onset of schizophrenia rather than an associated behaviour once the disorder develops. One way to do this is to use primary care electronic health records in which smoking data, psychiatric diagnoses and a number of other clinical and social variables are recorded over the life course. **These clinical data provide a unique resource to test the hypothesis that smoking initiation is associated with an increased risk of developing psychosis.**

### **Research question and hypotheses**

*Primary research question:*

Is smoking, as recorded in primary care data, related to **onset of** schizophrenia?

### **Hypotheses**

1. **Smoking initiation is associated with an increased risk of developing non-affective psychosis, including schizophrenia.**
2. **There is a dose response relationship between smoking and risk of schizophrenia**

## METHODS

### Data Source

We used data from The Health Improvement Network (THIN) primary care database, which contains data on over 12 million patients (including 3.7 million active patients) from over 500 general practices across the UK (<https://www.the-health-improvement-network.co.uk/#thin-research>). There are currently 8732 general practices in the UK. This database contains real time clinical data from GPs' electronic records which include symptoms, diagnosis and lifestyle behaviour such as smoking. The database is broadly representative of UK general practices in terms of patients' age and sex, practice size and geographical distribution. Furthermore, THIN practices contain a profile of patients that is representative of the UK population (9).

#### *Recording of smoking in GP records*

The U.K.'s quality and outcomes framework (QOF) for primary medical care was introduced in 2004 and rewards practices for meeting agreed performance targets(10). Indicators for recording smoking status and providing advice to smokers have featured since QOF was introduced. Thus, smoking status is relatively well recorded in UK electronic primary care records. In 2008-2009 smoking status was recorded for 84% of newly registered patients within 1 year of registration(11).

### Cohort

This study included data from all general practices that had contributed data to THIN in accordance with agreed standards. Patients were thus only eligible for entry after the date on which their practice had met the predefined standards for acceptable computer use and acceptable mortality recording (12, 13).

We identified all young people aged from 15 to 24 between 1 January 2004 and 31 December 2009. Individuals were eligible for entry to the cohort from the latest of: date of registration with the practice, dates of acceptable computer use and acceptable mortality recording by the practice, date of their fifteenth birthday, or 1 January 2004.

We followed up individuals until the earliest of: first diagnosis of schizophrenia (or psychosis), patient leaving the practice, practice leaving THIN, patient death, or 31 December 2014. Patients were aged between 15 and 34 years on leaving the study (some were in the cohort for less than a year), enabling us to observe them for up to 10 years, depending upon the date they entered the cohort.

Read codes are a list of clinical terms that have been used by the UK's National Health Service since 1985. They provide a standard vocabulary for clinicians to record patient findings and procedures. Our main exposures were smoking status and smoking intensity, which were defined according to Read codes in medical and additional health data records using established methods to compile the code lists(14). We are aware that individuals may have started smoking earlier, but smoking was defined as starting from the date of first record of smoking status. If someone had a record of smoking they were defined as a smoker; if they had a record of being a non smoker they were defined as a non smoker. People whose status changed from never smoker to smoker during the observation period were classed as smoking initiators. Those who had no record indicating that they were a smoker and no record indicating that they were a non smoker were excluded from the analysis. Patients were also excluded if they had a record of being an ex-smoker before being a non-smoker or a first smoking record, as this indicated misclassification of smoking status. An indicator of smoking intensity (never, light, moderate, heavy) was used based on highest recorded smoking intake during the first 12 months of the observation period.

Smoking intensity was classified by GPs as light (1-9 cigarettes per day), moderate (10-19 cigarettes per day), heavy (20-39 cigarettes per day) or very heavy (more than 40 cigarettes per day) when entering data into the system. Our primary outcome was schizophrenia and our secondary outcome was non-affective psychosis including schizophrenia. Clinicians may be reluctant to give a diagnosis of schizophrenia at first presentation and use a wider or less sensitive term. Thus, we also undertook a secondary analysis in which we included non-affective psychosis in addition to schizophrenia. The disorders were defined according to Read codes indicating a new diagnosis. The code list for a diagnosis of schizophrenia and non-affective psychosis was reviewed by two clinicians to ensure that affective and organic psychotic disorders were excluded. This list is available from the authors. Individuals who only had Read code records which were indicative of pre-existing schizophrenia or repeated episodes (such as chronic paranoid schizophrenia) were excluded. We likewise excluded individuals who had a record of a schizophrenia diagnosis or other non-affective psychotic disorder before their first smoking status record or up to 12 months afterwards. We chose 12 months as the entry of a diagnosis of SCZ often occurs after several months of dysfunction. This helped us protect against reverse causality, because we used smoking records from this period to define smoking intensity. For the main analyses, we therefore started the follow up period at 12 months after entry to the study (see analysis below).

The association between smoking and schizophrenia may potentially be confounded by social deprivation and recreational drug use. The Townsend score has been used for decades in the UK as a measure of social deprivation(15). It is derived from variables such as car ownership, owner occupier housing and overcrowded housing. We defined social deprivation in quintiles of the Townsend score for an individual up to 12 months after entry to the cohort. We excluded any GP practices where Townsend scores were poorly recorded.

We defined a binary indicator of history of recreational drug use on the basis of Read codes and once again, we used all records up to 12 months after entry to the cohort in generating this variable.

The study design is illustrated in Figure 1.

### **Analysis**

Binary indicator variables were created to identify patients who had ever had a diagnosis of (1) schizophrenia, or more broadly (2) non-affective psychosis (including schizophrenia).

The earliest recorded date of either schizophrenia or psychotic disorder was taken as the event date for both schizophrenia and psychosis. Incident rates of schizophrenia and non-affective psychosis (including schizophrenia) were estimated as the number of new diagnoses per 100,000 person years at risk (PYAR) from 12 months after entry to the study until end of follow up (defined as the earliest of: date of first diagnosis, date of death, date of patient leaving the practice, date of last data collection from practice or 31 December 2014).

We report unadjusted incidence rates by smoking status (non smoker; smoking initiator; smoker) and by smoking intensity (non, light, moderate and heavy smokers, smoking respectively 0, 1-9, 10-19 and 20 or more cigarettes per day).

Associations between smoking and incidence of schizophrenia and any non-affective psychosis were examined using multilevel Poisson regression models with (log) age as an offset. A random intercept of practice was specified to account for correlations between patients within practices. Results are reported as incident rate ratios (IRR). Similar models were used to examine the association between smoking intensity and incidence of schizophrenia and any non-affective psychotic disorder. In addition, we report IRR adjusted for quintiles of Townsend deprivation score.

Smoking intensity was defined using records up to 12 months after entry to the study. Individuals who were lost to follow up or with a first diagnosis during this period were excluded from the analysis. We did not perform any formal tests for a dose response effect, but rather inferred this relationship from the increasing magnitude of the estimates over the categories of increasing smoking and smoking intensity. A further analysis investigated the sensitivity of results to different methods of defining intensity by using smoking records up to 24 months after patient entry. This was conducted in order to obtain more and better information on smoking intensity.

We explored the possibility that confounding by cannabis use could explain associations between smoking and schizophrenia using the rule out method described by Schneeweiss (2006) (16) (Figure 2). This is a graphical method that quantifies the strength of confounding (i.e. the association between smoking and cannabis use AND the association between cannabis use and schizophrenia) needed to rule out an observed association. The graphs were informed by data on the prevalence of smoking and prevalence of cannabis as well as the IRR and 95%CI for the associations between smoking and schizophrenia. We did this using a combination of results from our study and figures reported in the literature. Thus, we estimated prevalence of smoking from our own data of 31.1% for men and 27.8% for women, and an estimated prevalence of cannabis use of 20.4% for men and 12.8% for women from the 2016-2017 crime survey for England and Wales which is representative of the UK population(17). These figures were appropriate, as the THIN population is representative of the UK population. The resulting curves divide the graph area in two. The area to the right and above each curve represents when an apparent association could be explained by confounding, i.e. combinations of associations with cannabis use (the potential confounder) which would reduce the IRR or the lower bound of the 95% CI to 1. In contrast,

the area to the left and below each curve represents the situations when association between smoking and schizophrenia/psychosis cannot be ruled out by residual confounding.

In Figure 2, we produced separate graphs for men and women and for incidence of schizophrenia and non-affective psychosis and interpreted them in conjunction with associations with cannabis use reported in the literature.

The THIN scheme for obtaining and providing anonymous patient data to researchers was approved by the NHS South-East Multicentre Research Ethics Committee in 2002. Approval for this study was obtained from the Scientific Review Committee (protocol reference number 16THIN041)

## RESULTS

### *Identification of cohort*

Initially, data on 1,056,176 patients were extracted. Smoking records from within the first 12 months after entry were used to establish smoking status and intensity. The observation period for incidence of schizophrenia or non-affective psychotic disorder thus started 12 months after entry to the study. 907,586 patients were retained in the analysis data set after excluding those who were lost to follow up before the start of the observation period (n=104,937), those with a first diagnosis earlier than 1 year after the date of first record of smoking (n=392), those missing a Townsend deprivation score (n=42,990) and patients whose smoking status could not be ascertained from Read codes (n=271).

### *Description of cohort*

*Diagnosis:* Of the 907,586 included in the analysis 536 (0.06%) received a diagnosis of schizophrenia and 1,582 (0.17%) received a diagnosis of non-affective psychotic disorder

(including schizophrenia). Overall, 258,170 (28.5%) of the cohort were aged 15 years on entry to the study, 267,771 (29.5%) aged 16 to 20 years, 307,256 (33.9%) aged 21 to 25 years and 74,389 (8.2%) aged 26 to 30 years. There were slightly more women (52.6%) than men (47.4%). The mean age of first diagnosis with a non-affective psychotic disorder (including schizophrenia) was 24.0 years (SD: 4.1) in males and 24.5 years (SD: 4.5) in females. The median length of follow up was 6.0 years (Inter Quartile Range: 3.4 to 9.0 years).

*Smoking:* 570,457 (62.9%) were classified as non-smokers throughout the whole observation period, 266,396 (29.4%) had started smoking prior to the observation period (smokers) and 70,733 (7.8%) started smoking during follow up (smoking initiators). The mean age of first record of smoking was 21.1 years (SD: 3.8) in males and 20.0 years (SD: 3.6) in females.

Based on smoking records up to 12 months after entry to the study, 465,192 (51.3%) patients were classified as never smokers, 71,503 (7.9%) as light ( $\leq 9$  cigarettes per day), 86,473 (9.5%) as moderate (10-19 cigarettes per day) and 31,930 (3.5%) as heavy smokers ( $\geq 20$  cigarettes per day). It was not possible to determine the level of smoking for 252,488 young people (27.8%) and these individuals were therefore not included in analyses of smoking intensity.

Using smoking records up to 24 months after entry to the study, 520,421 (57.32%) of patients were classified as never smokers, 80,332 (8.85%) as light, 100,631 (11.09%) as moderate and 39,494 (4.35%) as heavy smokers. It was not possible to determine level of smoking for 166,888 (18.39%) of patients.

*Recreational drug use:* In total, 33,512 (3.7%) individuals had a record indicating a history of recreational drug use. This is much lower than other estimates of recreational drug use

made in other epidemiological studies suggesting that many individuals who had used recreational drugs did not have a record of this in their primary care data and we therefore decided not to adjust for this covariate in our analysis.

*Incidence rates by smoking status and intensity*

*Incidence rates for schizophrenia:* Among men, the incidence rates for schizophrenia per 100,000 PYAR were higher in smoking initiators than in non-smokers (adjusted IRR: 1.94; 95% CI: 1.29 to 2.91) and higher still in smokers (adjusted IRR: 3.32; 95% CI: 2.67 to 4.14).

There was evidence of a dose response effect for amount of smoking, with an effect size of markedly greater magnitude in heavy smokers (adjusted IRR: 5.54; 95% CI: 3.84 to 7.99) than in moderate smokers (adjusted IRR: 3.80; 95% CI: 2.74 to 5.27). Adjustment for deprivation reduced the magnitude of the IRRs very slightly, suggesting only slight confounding (Table 1).

Among women, there was evidence of an elevated incidence rate of schizophrenia in smokers compared to non-smokers (adjusted IRR: 1.50; 95% CI: 1.06 to 2.12), but not in individuals who started smoking during the observation period. Similarly, only heavy smokers showed evidence of elevated rates of schizophrenia incidence compared to non-smokers (adjusted IRR: 3.74; 95% CI: 2.10 to 6.64). Again our adjustments had little impact on the result (Table 1).

*Incidence rates for non-affective psychosis (including schizophrenia):* Among men, psychosis incidence rates were higher in smoking initiators than in non-smokers (adjusted IRR: 2.65; 95% CI: 2.15 to 3.27) and higher still in smokers (adjusted IRR: 2.96; 95% CI: 2.59 to 3.38). There was evidence of a dose response effect for amount of smoking, with effect sizes of increasing magnitude in light (adjusted IRR: 2.72; 95% CI: 2.18 to 3.39), moderate (adjusted

IRR: 3.22; 95% CI: 2.67 to 3.88) and heavy smokers (adjusted IRR: 4.07; 95% CI: 3.26 to 5.09) relative to non-smokers (Table 2).

Among women, psychosis incidence rates were higher in smoking initiators than in non-smokers (adjusted IRR: 1.90; 95% CI: 1.40 to 2.56) and also in smokers (adjusted IRR: 2.13; 95% CI: 1.76 to 2.57). There was likewise evidence of a dose response effect for amount of smoking, with effect sizes of increasing magnitude in light (adjusted IRR: 2.20; 95% CI: 1.65 to 2.91), moderate (adjusted IRR: 2.31; 95% CI: 1.78 to 3.01) and heavy smokers (adjusted IRR: 4.10; 95% CI: 2.97 to 5.66) relative to non-smokers (Table 2).

*Sensitivity analysis:* The impact of smoking intensity was also examined based on smoking records up to 24 months after the start of the observation period (Table 3). This was to allow for a longer period of prodromal symptoms as might be expected before an entry of a diagnosis of schizophrenia or non-affective psychosis was made in the records. Higher smoking intensity in the 24 months before diagnosis is strongly linked in both sexes to an increased likelihood of either schizophrenia or the more inclusive category of non-affective psychosis.

*Sensitivity analysis examining confounding by cannabis use:* Based on the information available from our own data and an estimated prevalence of cannabis use in the population, we were able to quantify the likelihood of confounding by cannabis use. The graphs for men in Figure 2 suggest that unmeasured confounding needs to be substantial to rule out the observed association between smoking and schizophrenia/psychosis. For example, this would be the case if male smokers were 15 times more likely to also use cannabis than non-smokers AND cannabis users were 10 times more likely to develop schizophrenia. Using evidence from the literature, the point represented by an OR of 3.90 for schizophrenia in cannabis users compared to non-users (18) and an OR of 4.13 for cannabis use in smokers

compared to non-smokers (19), falls below and to the left both of the curve for the IRR and the lower bound of its 95% CI. This suggests that cannabis use cannot fully explain the association we found between smoking and schizophrenia. The same is true for the relationship between smoking and psychosis in both men and women. For schizophrenia in women (Figure 2), the point represented by these two ORs falls below the curve for the IRR, but above the curve for its 95% CI, indicating that the data are compatible with confounding by cannabis use.

## DISCUSSION

### *Findings*

Our results show associations between cigarette smoking and incidence of schizophrenia, as well as non-affective psychoses including schizophrenia, after adjustment for social deprivation. In men, incidence of schizophrenia was nearly twice as high in smoking initiators as non-smokers, and three times higher in pre-existing smokers, with a dose response effect of amount of smoking. This pattern was similar, but less pronounced, in women with evidence for a higher incidence rate in smokers compared to non-smokers, but not for smoking initiators.

In men, the findings for non-affective psychosis (including schizophrenia) were similar to those for schizophrenia. However, for women the pattern was more marked than in schizophrenia in that the association held for smoking initiators and smokers, and a dose response relationship occurred across all three categories of smoker. The greater magnitude of the association / dose response effect in non-affective psychosis in women may have occurred because of the low numbers of women with a diagnosis of schizophrenia, resulting in lower power to detect an association in that comparison. The gender differences observed may be an artefact of the generally later onset of schizophrenia in women and the cohort age group.

Our sensitivity analysis revealed that whether smoking intensity was defined using smoking records up to 12 or 24 months after the start of the observation period had little impact on the findings.

Our second sensitivity analysis using the rule out method suggests that it is unlikely that cannabis use can solely explain the association we found between smoking and schizophrenia in men, or between smoking and psychosis in either sex. It is possible that

cannabis use could confound the association between smoking and schizophrenia in women, however very few women were diagnosed with schizophrenia and this rather than a true association may explain this finding.

#### *The context of other research*

Our incidence rates for schizophrenia are comparable to the recent overall figure (also from the THIN database) of 9.2 per 100,000 PYAR for people aged 16-65 in the UK(20). Our findings echo the results of other recent studies in this area. In a recent meta-analysis of observational research, Gurillo et al reported a relative risk of new psychotic disorders in daily smokers versus non-smokers of 2.18 (95% CI 1.23–3.85)(4). Daily smokers developed psychotic illness at an earlier age than non-smokers (weighted mean difference –1.04 years, CI –1.82 to –0.26). There is also a suggestion that people with psychosis may start smoking at an earlier age than healthy controls (–0.44 years, CI –1.21 to 0.34). An association between early age of initiation of smoking and non-affective psychosis has also been reported in an Australian birth cohort study(21). Similar (albeit small) associations have been found between use of nicotine in smokeless form and onset of schizophrenia(22). Although there are cultural differences in that prevalence of smoking in patients with schizophrenia in non-western countries is lower (at least in women)(23-25), a higher risk in psychotic people persists. In a review of Chinese data(24), prevalence of smoking in women with schizophrenia was 4.3% compared with a general population rate of 3.4%. Although not reported in the publication, this amounts to an odds ratio (OR) of 1.28 for smoking in women with schizophrenia.

#### **Strengths and** *Limitations*

This observational study has several strengths in taking account of time from exposure to outcome, magnitude and direction of the association, biological gradient (intensity of

smoking), consistency, and specificity. Most important of all our results are generalisable as the analysis was conducted in a very large sample that was representative of the UK population and was longitudinal in nature. Nevertheless, there are limited demographic variables in THIN for which adjustment is possible and we cannot rule out residual confounding. In particular, we cannot be certain whether smoking patterns disguised a potential effect from recreational drugs, especially cannabis(26). The prevalence of data on recreational drug use was considered too low to be plausible in this age group. The 2016-2017 crime survey for England and Wales found that 19% of young people aged 16 to 24 had taken a drug in the previous year (and the figure is even higher a decade earlier)(27). Thus, primary care records have failed to identify 15% of the young people in our cohort, who have probably used an illegal drug.

We are conscious that we cannot rule out smoking initiation as a part of the prodrome of psychosis, given that prodromal symptoms (social isolation, academic difficulties, odd behaviours, and symptoms of depression or anxiety) may precede the appearance of frank schizophrenia by an average of 23 months(28). Our stipulation of at least a year between entry to the cohort and onset of schizophrenia, however, may go some way towards excluding this possibility. Our mean age of 'first smoking' may be slightly higher than other estimates, although it depends on how this is defined. Recent estimates suggest the most common age of *first trying* a cigarette is 10–15 years(29), while take up of regular smoking usually continues up to the early 20s(30). If participants with multiple smoking records had a record of being a smoker followed by a record of being a non smoker or an ex smoker, we inferred that they had given up smoking. However, individuals may have stopped smoking but lacked any record of having done so (i.e. becoming an ex or non smoker). Finally, smoking was not necessarily recorded during every appointment and, even if it had been,

may not have captured every change in an individual's smoking history. Thus, a degree of misclassification of smoking status and intensity is likely in our dataset.

### *Genetic factors*

There is a debate about possible genetic explanations for our observations. For example, it has been suggested that nicotine may overcome an observed dysregulation in schizophrenia in the expression of genes that encode the nicotinic acetylcholine receptor ( $\alpha 7$ ) (31, 32). However, an unbalanced gene expression could precede the onset of psychotic symptoms, with nicotine abuse and dependence merely acting as indicators of future psychosis (33). Another polymorphism (*rs1051730*) in the nicotinic acetylcholine receptor gene cluster *CHRNA5-A3-B4* has been reported to be associated with schizophrenia and with smoking intensity, but not with smoking initiation or with depression. A further study conducted in Iceland also supports the notion of pleiotropy for the observed comorbidity of addiction (including nicotine) and psychosis (34). In another approach to the study of possible causal effects of smoking on schizophrenia, Gage and colleagues reported little evidence of a causal association between smoking initiation and schizophrenia in a study using Mendelian randomisation (35). However, they could not exclude the possibility of a causal effect of heavier, lifetime exposure, rather than initiation, on the disorder. A more recent study that also used Mendelian randomisation in the UK Biobank data has indicated that the association between smoking and schizophrenia may be due in part to a causal effect of smoking (36).

### *Interpretation*

The magnitude and strength of the association between smoking and incidence of schizophrenia supports suggestions that smoking involves a risk through as yet unknown

pathways, or that it is an addictive condition which shares genetic risk with schizophrenia and other non- psychotic psychoses (34).

**ACKNOWLEDGEMENTS**

MK and RJ acknowledge the support of the UCLH NIHR Biomedical Research Centre.

**CONFLICT OF INTEREST**

There are no conflicts of interest for any of the authors

**FUNDING**

No independent funding was obtained to undertake this study.

## REFERENCES

1. de Leon J, Diaz FJ. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophrenia Research*. 2005;76(2–3):135-57.
2. Winterer G. Why do patients with schizophrenia smoke? *Current Opinion in Psychiatry*. 2010;23(2):112-9.
3. Whitton A, Green A, Pizzagalli D, Williams J, Brunette M. S244. Potent Dopamine D2 Receptor Antagonists Block the Reward-Enhancing Effects of Nicotine in Smokers With Schizophrenia. *Biological Psychiatry*. 2018;83(9, Supplement):S443.
4. Gurillo P, Jauhar S, Murray RM, MacCabe JH. Does tobacco use cause psychosis? Systematic review and meta-analysis. *The Lancet Psychiatry*. 2015;2(8):718-25.
5. Large M, MacCabe JH. Tobacco and psychosis: Not quite a smoking gun. *Australian and New Zealand Journal of Psychiatry*. 2015;49(10):859-60.
6. Arseneault L, Cannon M, Wittton J, Murray RM. Causal association between cannabis and psychosis: examination of the evidence. *The British Journal of Psychiatry*. 2004;184(2):110-7.
7. Gage SH, Hickman M, Zammit S. Association Between Cannabis and Psychosis: Epidemiologic Evidence. *Biological Psychiatry*. 2016;79(7):549-56.
8. Murray RM, Bhavsar V, Tripoli G, Howes O. 30 Years on: How the Neurodevelopmental Hypothesis of Schizophrenia Morphed into the Developmental Risk Factor Model of Psychosis. *Schizophrenia Bulletin*. 2017:sbx121-sbx.
9. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. *Journal of Innovation in Health Informatics*. 2011;19(4):251-5.
10. Taggar JS, Coleman T, Lewis S, Szatkowski L. The impact of the Quality and Outcomes Framework (QOF) on the recording of smoking targets in primary care medical records: cross-sectional analyses from The Health Improvement Network (THIN) database. *BMC Public Health*. 2012;12(1):1-11.
11. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, White IR, et al. Smoker, ex-smoker or non-smoker? The validity of routinely recorded smoking status in UK primary care: a cross-sectional study. *BMJ Open*. 2014;4(4).
12. Horsfall L, Walters K, Petersen I. Identifying periods of acceptable computer usage in primary care research databases. *Pharmacoepidemiology and Drug Safety*. 2013;22(1):64-9.
13. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. *Pharmacoepidemiology and Drug Safety*. 2009;18(1):76-83.
14. Dave S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. *PharmacoepidemiolDrug Saf*. 2009;18(8):704-7.
15. Dolan SA, Jarman B, Bajekal M, Davies PM, Hart D. Measuring disadvantage: changes in the underprivileged area, Townsend, and Carstairs scores 1981-91. *Journal of Epidemiology and Community Health*. 1995;49(Suppl 2):S30-S3.
16. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and Drug Safety*. 2006;15(5):291-303.
17. Office H. Drug misuse: findings from the 2017 to 2018 CSEW. Appendix table 3.02. 2018.
18. Marconi A, Di Forti M, Lewis CM, Murray RM, Vassos E. Meta-analysis of the Association Between the Level of Cannabis Use and Risk of Psychosis. *Schizophrenia Bulletin*. 2016;42(5):1262-9.
19. Fix BV, Smith D, O'Connor R, Heckman BW, Willemsen MC, Cummings M, et al. Cannabis use among a nationally representative cross-sectional sample of smokers and non-smokers in the

- Netherlands: results from the 2015 ITC Netherlands Gold Magic Survey. *BMJ Open*. 2019;9(3):E024497.
20. Hardoon S, Hayes JF, Blackburn R, Petersen I, Walters K, Nazareth I, et al. Recording of severe mental illness in United Kingdom primary care, 2000-2010. *PloS one*. 2013;8(12):e82365-e.
  21. McGrath JJ, Alati R, Clavarino A, Williams GM, Bor W, Najman JM, et al. Age at first tobacco use and risk of subsequent psychosis-related outcomes: A birth cohort study. *Australian and New Zealand Journal of Psychiatry*. 2015.
  22. Munafò MR, Larsson Lönn S, Sundquist J, Sundquist K, Kendler K. Snus use and risk of schizophrenia and non-affective psychosis. *Drug and Alcohol Dependence*. 2016;164:179-82.
  23. Aguocha CM, Aguocha JK, Igwe M, Uwakwe RU, Onyeama GM. Prevalence and correlates of cigarette smoking among patients with schizophrenia in southeast Nigeria. *Acta Psychiatrica Scandinavica*. 2015;131(3):206-12.
  24. Cao X-L, Li Y, Zhong B-L, Chiu HFK, Ungvari GS, Ng CH, et al. Current cigarette smoking in Chinese female patients with schizophrenia: A meta-analysis. *Psychiatry Research*. 2016;235:203-5.
  25. Li Y, Hou C-L, Ma X-R, Zang Y, Jia F-J, Zhong B-L, et al. Smoking and its associations with sociodemographic and clinical characteristics and quality of life in patients with schizophrenia treated in primary care in China. *General Hospital Psychiatry*. 2016;38:79-83.
  26. Fergusson DM, Hall W, Boden JM, Horwood LJ. Rethinking cigarette smoking, cannabis use, and psychosis. *The Lancet Psychiatry*. 2015;2(7):581-2.
  27. Broadfield D. Drug Misuse: Findings from the 2016/2017 Crime Survey for England and Wales. In: Office H, editor. *Statistical Bulletin 11/17* ed. UK: Home Office; 2017. p. 2.
  28. Kafali HY, Bildik T, Bora E, Yuncu Z, Erermis HS. Distinguishing prodromal stage of bipolar disorder and early onset schizophrenia spectrum disorders during adolescence. *Psychiatry Research*. 2019;275:315-25.
  29. West R. Tobacco smoking: Health impact, prevalence, correlates and interventions. *Psychology & Health*. 2017;32(8):1018-36.
  30. Dierker L, He J, Kalaydjian A, Swendsen J, Degenhardt L, Glantz M, et al. The importance of timing of transitions for risk of regular smoking and nicotine dependence. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine*. 2008;36(1):87-92.
  31. Kunii Y, Zhang W, Xu Q, Hyde TM, McFadden W, Shin JH, et al. *CHRNA7* and *CHRFAM7A* mRNAs: Co-Localized and Their Expression Levels Altered in the Postmortem Dorsolateral Prefrontal Cortex in Major Psychiatric Disorders. *American Journal of Psychiatry*. 2015;172(11):1122-30.
  32. Liu X-a, Kenny PJ.  $\alpha 5$  nicotinic receptors link smoking to schizophrenia. *Nature Medicine*. 2017;23:277.
  33. Akbarian S, Kundakovic M. *CHRNA7* and *CHRFAM7A*: Psychosis and Smoking? Blame the Neighbors! *American Journal of Psychiatry*. 2015;172(11):1054-6.
  34. Reginsson GW, Ingason A, Euesden J, Bjornsdottir G, Olafsson S, Sigurdsson E, et al. Polygenic risk scores for schizophrenia and bipolar disorder associate with addiction. *Addiction Biology*. 2018;23(1):485-92.
  35. Gage SH, Jones HJ, Taylor AE, Burgess S, Zammit S, Munafò MR. Investigating causality in associations between smoking initiation and schizophrenia using Mendelian randomization. *Scientific Reports*. 2017;7:40653.
  36. Wootton RE, Richmond RC, Stuijzand BG, Lawn RB, Sallis HM, Taylor GMJ, et al. Causal effects of lifetime smoking on risk for depression and schizophrenia: Evidence from a Mendelian randomisation study. *bioRxiv*. 2018:381301.

**Table 1: Incidence of schizophrenia by smoking status and intensity**

	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) p value	Adjusted IRR (95% CI) p value	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) p value	Adjusted IRR (95% CI) p value
<b>Smoking status</b>										
	<b>Men (N = 429,922)</b>					<b>Women (N = 477,664)</b>				
Non smoker	123	14.36	8.6 (7.2 to 10.2)	1	1	76	14.20	5.4 (4.3 to 6.7)	1	1
Smoking initiator	29	2.13	13.6 (9.5 to 19.6)	2.00 (1.33 to 2.99)	1.94 (1.29 to 2.91)	12	2.35	5.1 (2.9 to 9.0)	1.28 (0.69 to 2.35)	1.25 (0.68 to 2.29)
Smoker	239	7.57	31.6 (27.8 to 35.8)	3.56 (2.87 to 4.43)	3.32 (2.67 to 4.14)	57	6.59	8.6 (6.7 to 11.2)	1.65 (1.17 to 2.32)	1.50 (1.06 to 2.12)
				p<0.0001	p<0.0001				p=0.0172	p=0.0711
<b>Smoking intensity *</b>										
	<b>Men (N = 279,007)</b>					<b>Women (N = 376,091)</b>				
Non smoker	81	9.63	8.4 (6.8 to 10.5)	1	1	56	11.69	4.8 (3.7 to 6.2)	1	1
Light smoker	49	1.30	37.6 (28.4 to 49.8)	4.08 (2.86 to 5.82)	3.91 (2.74 to 5.58)	12	2.00	6.0 (3.4 to 10.6)	1.25 (0.67 to 2.33)	1.20 (0.64 to 2.23)
Moderate smoker	67	1.80	37.3 (29.4 to 47.4)	4.04 (2.92 to 5.59)	3.80 (2.74 to 5.27)	19	2.38	8.0 (5.1 to 12.5)	1.73 (1.03 to 2.91)	1.58 (0.94 to 2.67)
Heavy smoker	46	0.76	60.2 (45.1 to 80.3)	6.08 (4.23 to 8.74)	5.54 (3.84 to 7.99)	15	0.73	20.6 (12.4 to 34.2)	4.29 (2.42 to 7.59)	3.74 (2.10 to 6.64)
				p<0.0001	p<0.0001				p<0.0001	p<0.0001

**Notes:** N = number of incident cases, PYAR = person years at risk, IR = incident rate per 100,000 PYAR, IRR = incident rate ratio, CI = confidence interval. Adjusted IRR control for Townsend deprivation score. \* Smoking intensity was defined using records up to 12 months after entry to the study. Light smoker: <=9 cigarettes per day; moderate smoker: 10-19 cigarettes per day; heavy smoker: >=20 cigarettes per day.

**Table 2: Incidence of all non affective psychosis (including schizophrenia) by smoking status and intensity**

	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) <i>p value</i>	Adjusted IRR (95% CI) <i>p value</i>
<b>Smoking status</b>										
	<b>Men (N = 429,922)</b>					<b>Women (N = 477,664)</b>				
Non smoker	354	14.36	24.6 (22.2 to 27.4)	1	1	219	14.20	15.4 (13.5 to 17.6)	1	1
Smoking initiator	116	2.13	54.4 (45.4 to 65.3)	2.71 (2.20 to 3.35)	2.65 (2.15 to 3.27)	53	2.35	22.6 (17.3 to 29.6)	1.93 (1.43 to 2.61)	1.90 (1.40 to 2.56)
Smoker	613	7.57	81.0 (74.8 to 87.7)	3.15 (2.76 to 3.59)	2.96 (2.59 to 3.38)	227	6.59	34.4 (30.2 to 39.2)	2.27 (1.88 to 2.73)	2.13 (1.76 to 2.57)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> <0.0001	<i>p</i> <0.0001
<b>Smoking intensity *</b>										
	<b>Men (N = 279,007)</b>					<b>Women (N = 376,091)</b>				
Non smoker	269	9.63	27.9 (24.8 to 31.5)	1	1	173	11.69	14.8 (12.8 to 17.2)	1	1
Light smoker	113	1.30	86.7 (72.1 to 104.3)	2.82 (2.26 to 3.52)	2.72 (2.18 to 3.39)	67	2.00	33.5 (26.3 to 42.5)	2.26 (1.70 to 2.99)	2.20 (1.65 to 2.91)
Moderate smoker	190	1.80	105.8 (91.8 to 122.0)	3.41 (2.83 to 4.11)	3.22 (2.67 to 3.88)	84	2.38	35.3 (28.5 to 43.7)	2.45 (1.89 to 3.18)	2.31 (1.78 to 3.01)
Heavy smoker	114	0.76	149.1 (124.1 to 179.1)	4.45 (3.57 to 5.55)	4.07 (3.26 to 5.09)	49	0.73	67.3 (50.9 to 89.1)	4.47 (3.25 to 6.15)	4.10 (2.97 to 5.66)
				<i>p</i> <0.0001	<i>p</i> <0.0001				<i>p</i> <0.0001	<i>p</i> <0.0001

**Notes:** N = number of incident cases, PYAR = person years at risk, IR = incident rate per 100,000 PYAR, IRR = incident rate ratio, CI = confidence interval. Adjusted IRR control for Townsend deprivation score. \* Smoking intensity was defined using records up to 12 months after entry to the study. Light smoker: <=9 cigarettes per day; moderate smoker: 10-19 cigarettes per day; heavy smoker: >=20 cigarettes per day.

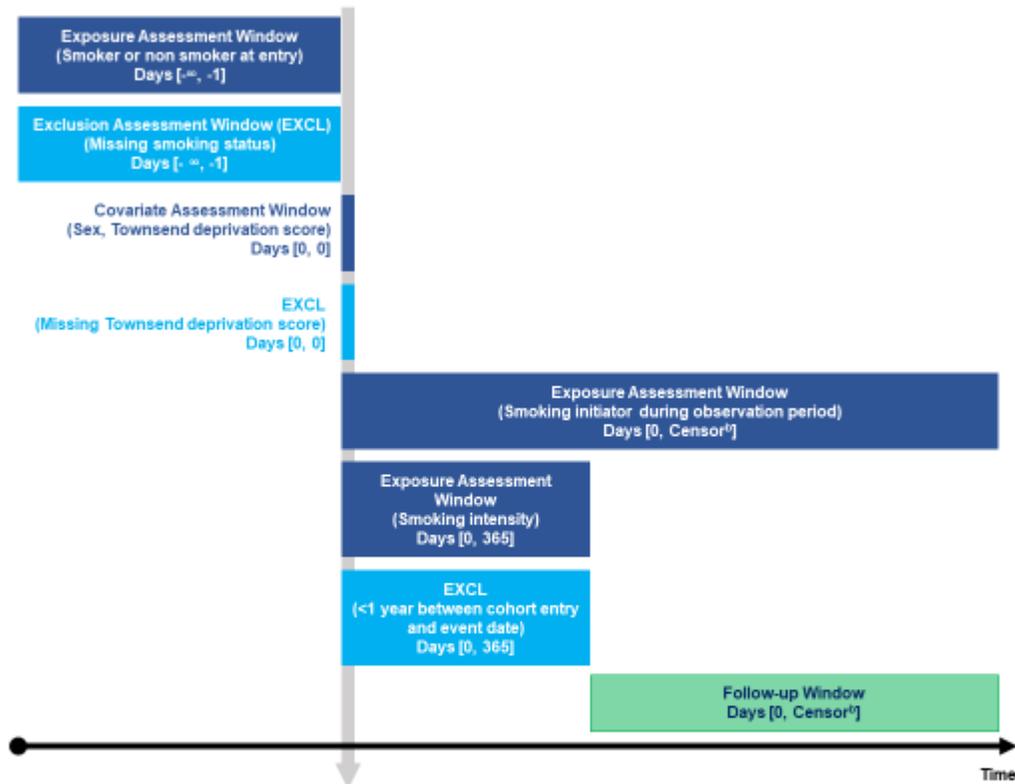
**Table 3: Association between smoking intensity and incidence of schizophrenia or non-affective psychosis based on smoking records up to 24 months after entry to the study**

	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) p value	Adjusted IRR (95% CI) p value	N	100,000 PYAR	IR (95% CI)	Unadjusted IRR (95% CI) p value	Adjusted IRR (95% CI) p value
<b>Schizophrenia</b>										
	<b>Men (N = 285,562)</b>					<b>Women (N = 355,139)</b>				
Non smoker	80	11.36	7.0 (5.7 to 8.8)	1	1	54	13.04	4.1 (3.2 to 5.4)	1	1
Light smoker	48	1.55	31.0 (23.3 to 41.1)	4.17 (2.91 to 5.96)	4.03 (2.81 to 5.77)	8	2.24	3.6 (1.8 to 7.1)	0.87 (0.41 to 1.82)	0.83 (0.40 to 1.75)
Moderate smoker	66	2.19	30.2 (23.7 to 38.4)	4.06 (2.93 to 5.62)	3.85 (2.78 to 5.35)	14	2.80	5.0 (3.0 to 8.4)	1.25 (0.69 to 2.25)	1.14 (0.63 to 2.05)
Heavy smoker	42	0.97	43.3 (32.0 to 58.6)	5.53 (3.80 to 8.04)	5.09 (3.49 to 7.43)	14	0.93	15.0 (8.9 to 25.3)	3.63 (2.02 to 6.54)	3.15 (1.74 to 5.71)
				p<0.0001	p<0.0001				p=0.0003	p=0.0013
<b>All non affective psychosis</b>										
	<b>Men (N = 285,562)</b>					<b>Women (N = 355,139)</b>				
Non smoker	266	11.36	23.4 (20.8 to 26.4)	1	1	161	13.04	12.3 (10.6 to 14.4)	1	1
Light smoker	106	1.55	68.4 (56.5 to 82.7)	2.77 (2.21 to 3.47)	2.67 (2.13 to 3.34)	50	2.24	22.3 (16.9 to 29.4)	1.81 (1.32 to 2.49)	1.76 (1.28 to 2.42)
Moderate smoker	196	2.19	89.6 (77.9 to 103.1)	3.58 (2.98 to 4.31)	3.39 (2.81 to 4.08)	72	2.80	25.7 (20.4 to 32.4)	2.14 (1.62 to 2.83)	2.01 (1.52 to 2.66)
Heavy smoker	115	0.97	118.6 (98.8 to 142.3)	4.46 (3.58 to 5.55)	4.09 (3.28 to 5.11)	45	0.93	48.2 (36.0 to 64.5)	3.88 (2.78 to 5.40)	3.53 (2.53 to 4.93)
				p<0.0001	p<0.0001				p<0.0001	p<0.0001

**Notes:** N = number of incident cases, PYAR = person years at risk, IR = incident rate per 100,000 PYAR, IRR = incident rate ratio, CI = confidence interval. Adjusted IRR control for Townsend deprivation score. Light smoker: <=9 cigarettes per day; moderate smoker: 10-19 cigarettes per day; heavy smoker: >=20 cigarettes per day.

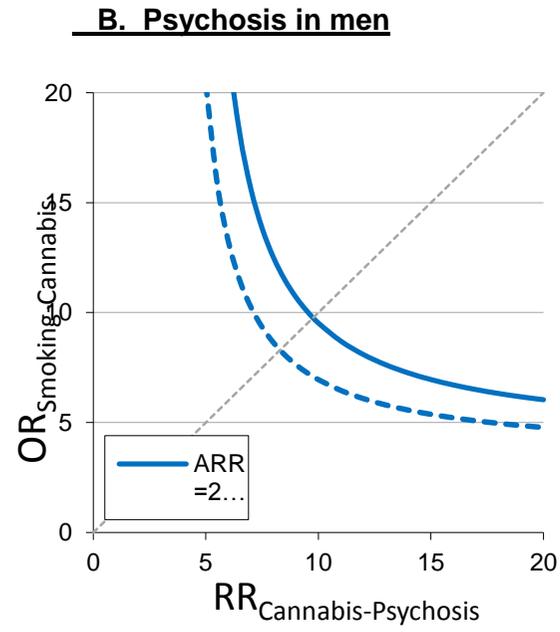
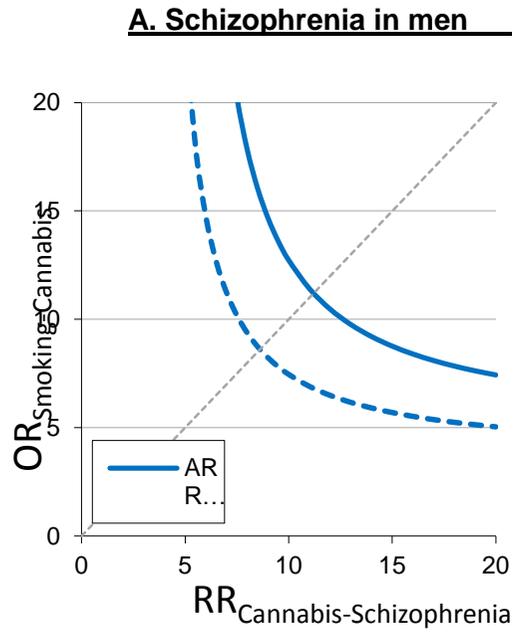
Figure 1 Study Design

Cohort Entry Date (Age 15 to 24 years between 1 January 2004 and 31 December 2009<sup>a</sup>) Day 0

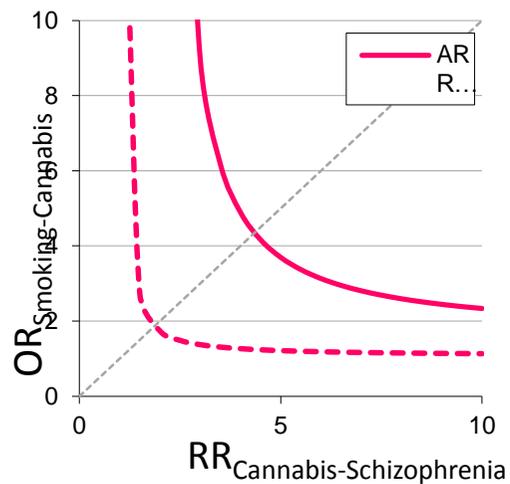


- Individuals were eligible for entry to the cohort from the latest of: date of registration with the practice, dates of acceptable computer use and acceptable mortality recording by the practice, date of their fifteenth birthday, or 1 January 2004.
- Individuals were followed up until the earliest of: first diagnosis of schizophrenia (or psychosis), patient leaving the practice, practice leaving THIN, patient death, or 31 December 2014.

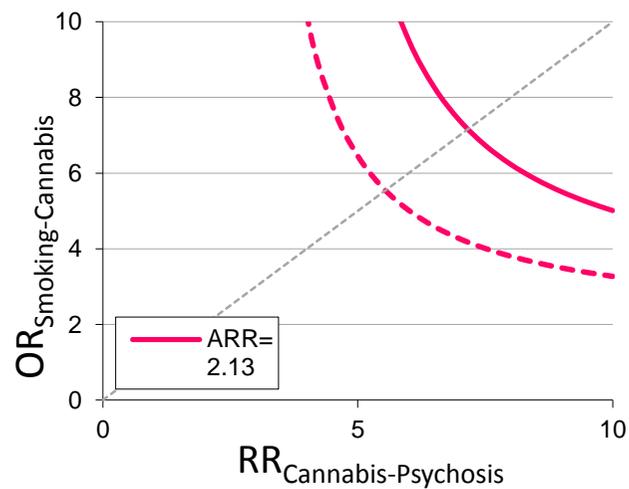
**Figure 2:** Cannabis use as a potential confounder of the apparent association between smoking and schizophrenia and non-affective psychosis



### C. Schizophrenia in women

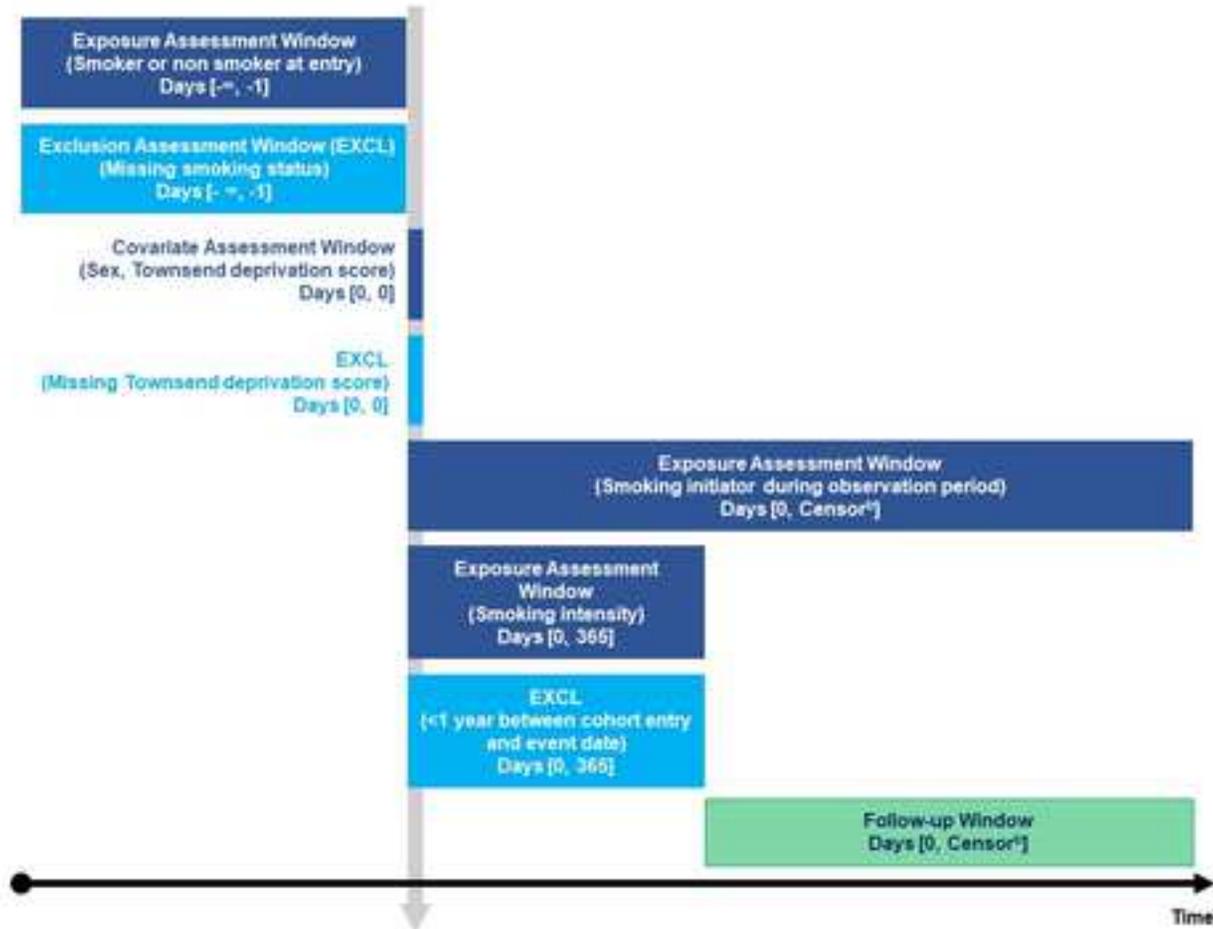


### D. Psychosis in women



**Figure note/caption:** Solid lines represent the adjusted incident rate ratio (IRR) for incidence of schizophrenia or psychosis respectively in smokers compared to non smokers. Dashed lines represent the lower bound of the 95% confidence interval (CI) for the IRR. The areas of the graphs above and to the right of these curves indicate combinations of associations with cannabis use (the potential confounder) which would reduce the IRR or the lower bound of the 95% CI to 1, indicating that the confounding by cannabis use eliminates the association between smoking and schizophrenia or psychosis. Conversely, areas below and to the left of the curves indicate combinations which may result in partial confounding, but where cannabis use cannot fully explain the observed association.

Figure 1 Study Design

Cohort Entry Date (Age 15 to 24 years between 1 January 2004 and 31 December 2009<sup>a</sup>) Day 0

a. Individuals were eligible for entry to the cohort from the latest of: date of registration with the practice, dates of acceptable computer use and acceptable mortality recording by the practice, date of their fifteenth birthday, or 1 January 2004.

b. Individuals were followed up until the earliest of: first diagnosis of schizophrenia (or psychosis), patient leaving the practice, practice leaving THN, patient death, or 31 December 2014.

